Technical Note

Rapid Assembly of gRNA Arrays via Modular Cloning in Yeast

Nicholas S McCarty, William M Shaw, Tom Ellis, and Rodrigo Ledesma-Amaro

ACS Synth. Biol., **Just Accepted Manuscript** • Publication Date (Web): 02 Apr 2019

Rapid Assembly of gRNA Arrays via Modular Cloning in Yeast

Nicholas S. McCarty[1,2], William M. Shaw[2,3], Tom Ellis[2,3] & Rodrigo Ledesma-Amaro[2,3,4,*]

1 Bioengineering, California Institute of Technology, Pasadena, CA 91125, USA

2 Imperial College Centre for Synthetic Biology, Imperial College London, London, SW7 2AZ, UK

3 Department of Bioengineering, Imperial College London, London, SW7 2AZ, UK

4 Lead contact

*Correspondence: r.ledesma-amaro@imperial.ac.uk

ABSTRACT

CRISPR is a versatile technology for genomic editing and regulation, but the expression of multiple gRNAs in *S. cerevisiae* has thus far been limited. We present here a simple extension to the Yeast MoClo Toolkit, which enables the rapid assembly of gRNA arrays using a minimal set of parts. Using a dual-PCR, Type IIs restriction enzyme Golden Gate assembly approach, at least 12 gRNAs can be assembled and expressed from a single transcriptional unit. We demonstrate that these gRNA arrays can stably regulate gene expression in a synergistic manner via dCas9-mediated repression. This approach expands the number of gRNAs that can be expressed in this model organism and may enable the versatile editing or transcriptional regulation of a greater number of genes *in vivo*.

Keywords: yeast, MoClo toolkit, CRISPR, multiplex, assembly method, Golden Gate assembly

The expression of many gRNAs simultaneously enables multiplexed genome engineering and transcriptional modulation of genetic targets with CRISPR-Cas systems. There are numerous methods to accomplish this, such as by encoding each gRNA under the control of a separate RNA polymerase III (Pol III) promoter[1], cleaving gRNAs from a longer RNA sequence by encoding them in introns[2], by flanking gRNAs with self-processing ribozyme sequences[3], or via excision of gRNAs by other Cas family endonucleases[4–6]. This last approach, in which gRNAs are excised by an endonuclease, can be performed with Csy4, which processes pre-crRNA in native CRISPR systems[7]. By flanking gRNAs with Csy4 recognition sites, multiple crRNAs or

fused gRNAs can be efficiently processed from a single transcript in a variety of cell lines[5,6].

Numerous methods have also been reported to assemble gRNA arrays for *in vitro* and *in vivo* applications, typically by utilizing annealed oligonucleotides with modified Gibson or Golden Gate assembly approaches[6,8–12]. These studies have largely focused on mammalian cells but methods have recently emerged for gRNA multiplexing in yeast. Efforts to multiplex gRNAs in *S. cerevisiae* have been limited to 7 gRNAs if expressed from individual promoters[13] and 4 gRNAs if expressed and processed from a single array, either by using gRNA-tRNA arrays with Pol III promoters or by using Pol II promoters with the Csy4 enzyme to process the gRNA arrays.[5,14]

Here, we expand the widely-used Yeast MoClo Toolkit[1], a widely-used library of highly characterized parts designed for Golden Gate assembly and integration into *S. cerevisiae*, to enable the rapid, modular assembly and expression of gRNA arrays in this model organism. This method is a dual-PCR, Type IIs restriction enzyme Golden Gate approach that enables the construction of up to 12 gRNAs in a single array, each flanked by recognition sites for *Csy4*, in two days. This method uses one plasmid to produce gRNAs with desired spacer sequences and a second plasmid to assemble the final arrays **(Supplementary Figure S1)**. All other parts are available from the Yeast MoClo Toolkit.[1] This system will enhance the ease and utility of multiplexed gRNA applications in *S. cerevisiae.*

RESULTS

We build and express arrays of tandem gRNAs in *S. cerevisiae* by adding just two parts to the Yeast MoClo Toolkit: a Guide-Generating Vector (GGV) and a Destination Vector (DV) (**Figure 1a and Supplementary Figures S2, S3**). An initial round of PCR using the GGV as a template adds desired spacer sequences via primer overhangs. The GGV contains a *Csy4* recognition site, a selection/screening marker and a sequence encoding the gRNA scaffold of interest (which associates the gRNA with a chosen Cas endonuclease). Once the linear PCR products are ligated, the designed spacers are positioned next to the scaffold to produce functional gRNAs.

A second round of PCR adds Type IIs restriction enzyme recognition sites with unique, four-nucleotide overhangs via primer overhangs[15]. The generated, linear PCR fragments can then be assembled into the DV via Golden Gate assembly, which consists of a promoter of interest, followed by another screening/selection marker, a terminal copy of the *Csy4* recognition sequence, and the terminator of interest (**Figure 1a**). Target sequences of gRNAs in the array must be devoid of BsmBI, BsaI and NotI recognition sites. For this study, we use a DV with a start codon-deficient variant of P_{TDH3}, a Pol II promoter derived from the Yeast MoClo Toolkit[1]. The DV also contains designed XhoI and BglII sites after the promoter and before the final Csy4 site, respectively, which enables assembled gRNA arrays to be swapped between promoters and terminators from the Yeast MoClo Toolkit.

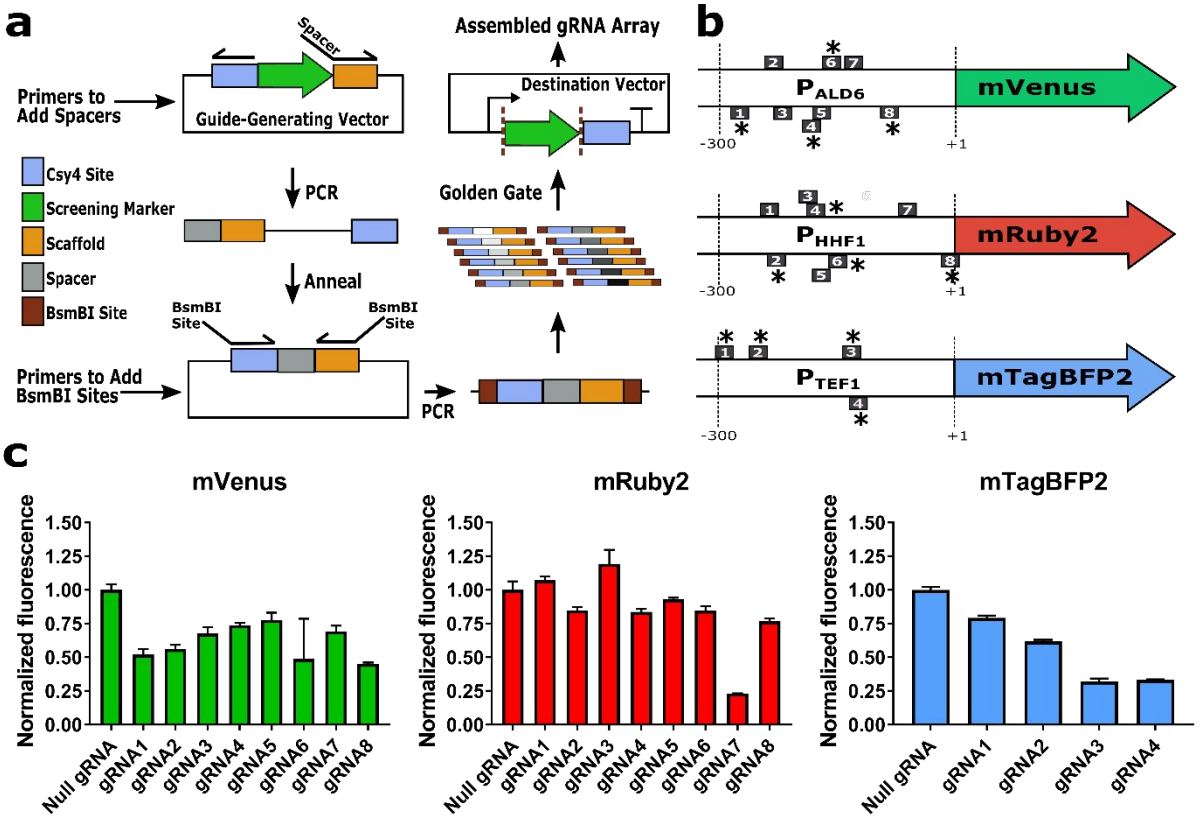


Figure 1: *Assembly of gRNA arrays and screening of gRNA candidates.* (a) Schematic overview of the system for assembling gRNA arrays. Primers are indicated by arrows, with slanted lines indicating primer overhangs. (b) Spatial positions of the spacer sequences tested on their corresponding genomic targets. Each gRNA contains a 20-nucleotide spacer sequence complementary to the P_{ALD6}, P_{HHF1} or P_{TEF1} promoter and is adjacent to an NGG PAM sequence. Numbers in the gray boxes indicate each unique gRNA tested and corresponds to the results plotted in panel (c). Numbers marked with an asterisk indicate gRNAs which were used in the construction of gRNA arrays. A minimal number of PAM sites

in the P_{TEF1} promoter limited the number of gRNAs which could be tested for mTagBFP2. (c) Normalized fluorescence for each individual gRNA tested. Each sample was normalized to the fluorescence measured for a yeast strain expressing a null spacer sequence (100% off-target score using Doench *et al.* algorithm)[16]. The fluorescence of yeast expressing the null gRNA was set to a value of 1.0 for each fluorescent reporter. All other groups were normalized to the null gRNA mean. Error bars represent the standard deviation between the n=3 replicates within each group.

To assess the efficiency of gRNA array assembly using this system, arrays containing between 3 to 12 gRNAs within a single transcriptional unit were constructed. Colonies were screened for the absence of the screening marker (GFP) (**Supplementary Figure S4**) and the insert length of each construct was evaluated by diagnostic restriction digest with *BsaI* (**Supplementary Figure S5**). Restriction digests indicated that 100% of colonies contained an array of the expected size (for 3 or 6 gRNAs), followed by a decrease in efficiency for 9 and 12 gRNA arrays (66.6% and 50%, respectively) (**Supplementary Figure S6**). Purified DNA isolated from colonies that yielded bands of the expected sizes from restriction digests were additionally sent for Sanger sequencing (n=3 each) for the 3, 6, 9 and 12 gRNA arrays. All samples were sequence-verified without any observed mutations (**Supplementary Methods**).

To demonstrate gRNA multiplexing capacity, an experiment was implemented that enabled us to test the transcriptional repression mediated by multiple gRNAs in budding yeast simultaneously using fluorescent reporters and flow cytometry. *S. cerevisiae* strain BY4741 was engineered to express three fluorescent reporters, each driven by a unique constitutive promoter: P_{TEF1}-mTagBFP2, P_{HHF1}-mRuby2, and P_{ALD6}-Venus. These promoters are constitutive, strong promoters from the Yeast MoClo Toolkit derived from genes in *S. cerevisiae* that are either nonessential or have paralogs in the genome. These promoter-reporter pairs were cloned into a single cassette and genome-integrated at the HO locus. A DNA construct expressing dCas9 (with N- and C-termini nuclear localization signals) and Csy4 endonuclease, under control of P_{PGK1} and P_{HHF2}, respectively, were also genome-integrated at the LEU2 locus[17].

Individual gRNAs were next designed to repress fluorescent protein expression by targeting each of the three constitutive promoters and their individual efficiencies were assessed. Spacer sequences to repress P_{TEF1}, P_{HHF1}, or P_{ALD6} were targeted to bind

between positions -300 and +1 bp upstream of the start codon of each fluorescent reporter (**Figure 1b**). The BY4741-derived strain expressing these fluorescent reporters, dCas9 and Csy4 was then transformed with individual gRNAs possessing the designed spacer sequences integrated at the URA3 locus. Changes in fluorescence of each reporter following integration of the gRNA were assessed by flow cytometry (**Figure 1c**).

Using these data, four gRNAs were selected for dCas9-mediated transcriptional repression for each of the fluorescent reporters. They were selected based on weak repression of fluorescent output, so that synergistic effects could be visualized when multiplexing many gRNAs to a promoter, and distributed targeting within the promoter regions so that gRNA recognition sites do not overlap, as we speculated that this would occlude our ability to observe synergistic dCas9-mediated repression of multiple targets simultaneously. gRNAs V1, V4, V6, V8 for Venus repression, gRNAs R2, R8, R6, R4 for mRuby2 repression, and gRNAs B1, B2, B3, B4 for mTagBFP2 repression (**Figure 1b**) were selected.

Arrays encoding 3, 6, 9, or 12 gRNAs within a single transcriptional unit were constructed. As arrays increase in size, more gRNAs are targeted to repress fluorescent protein expression; the 3-gRNA array targets one gRNA per promoter, whereas the 12-gRNA array targets four gRNAs to each promoter. The 3-gRNA array consists of V1, R2 and B1. The 6-gRNA array consists of V1, R2, B1, V4, R8 and B2. The 9-gRNA array consists of V1, R2, B1, V4, R8, B2, V6, R6 and B3. The 12-gRNA array consists of all twelve selected gRNAs. All arrays were genome-integrated at the URA3 locus (**Figure 2a**). We observed a slight growth defect in the strains expressing gRNA arrays compared to controls, potentially due to a side targeting of the three native promoters.[18]

Characterization of per cell fluorescence by flow cytometry revealed synergistic repression of all three fluorescent proteins when an increasing number of gRNAs were expressed from the arrays (**Figure 2b**). Co-expression of dCas9 with four gRNAs targeting each fluorescent reporter, in the case of the 12-gRNA array, showed a repression of 92% in the mVenus channel, 81% in the mRuby2 channel and 95% in the mTagBFP2 channel. These repression efficiencies are significantly greater than

with any of the individual gRNAs and indicated synergistic repression of three separate genes by a tandem array of up to 12 gRNAs using this multiplexed system.

Since homologous recombination in budding yeast is more active in regions containing repetitive DNA sequences[19], we also assessed the stability of these repetitive gRNA arrays over generations of growth by performing flow cytometry on yeast expressing each gRNA array after five consecutive, 1:100 dilutions with twelve hours of growth between dilutions; 72 hours of growth in total (**Supplementary Figure S7**). Flow cytometry after five consecutive dilutions indicated sustained, synergistic transcriptional repression of fluorescent protein expression. Colony PCR on the gRNA arrays at the URA3 locus confirmed stable genomic integration of all gRNA arrays after the dilution series (**Supplementary Figure S7**).

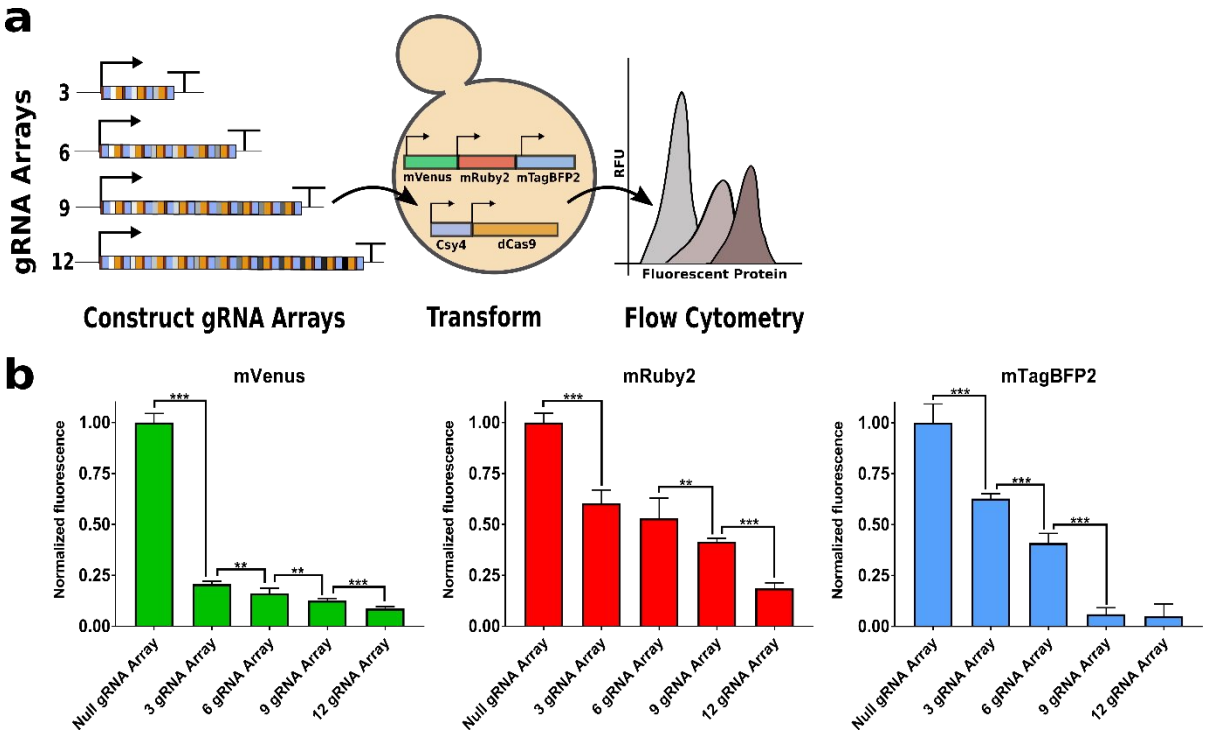


Figure 2: *Genome-integrated gRNA arrays in S. cerevisiae enable synergistic, dCas9-mediated repression of gene expression.* (a) General method to quantify the repression efficiency of tandem gRNA arrays in *S. cerevisiae*. gRNA arrays were driven by P_{TDH3} on a plasmid integrated at the URA3 locus. Arrays contain 3, 6, 9, or 12 gRNAs in tandem, with each successive increase in tandem length adding an additional gRNA targeting each of the promoters driving fluorescent protein expression. gRNA arrays were transformed into yeast expressing Csy4, dCas9 and Venus, mRuby2 and mTagBFP2 and fluorescence was measured by flow cytometry to quantify the repression efficiency of each gRNA array. (b) Normalized repression of fluorescence by multiplexed expression of gRNA arrays. All experimental samples were normalized to the fluorescence measured for this yeast strain also expressing an array with 3 'null' gRNAs. The '3 Null gRNAs' control was set to a value of 1.0 for all three fluorescence channels. All values plotted are averages from n=8 (3, 6, 9, 12

gRNA arrays) or n=4 (3 Null gRNAs) samples. Error bars represent one standard deviation from the geometric mean. Asterisks denote two-tail *p*-value as determined by two-sample *t*-test, with $^{**}p \leq 0.01$, and $^{***}p \leq 0.0001$.

This assembly method is rapid, inexpensive, and fully compatible with the existing Yeast MoClo Toolkit[1]. It can be used to express at least 12 gRNAs from one promoter, which may prove useful in multiplexed CRISPR applications, such as the transcriptional regulation of multiple genes in a pathway, or to construct more complex, gRNA-based logic circuits[13].

MATERIALS AND METHODS

Assembly and Validation of gRNA Arrays

A detailed protocol, including primer sequences and thermocycler settings, can be found in the Supporting Information (**Supplementary Tables S1-S13**). 5 ng of GGV template was PCR amplified with a phosphorylated reverse primer and a forward primer with 20 nucleotide overhang to add spacer sequences. PCR mixtures were then digested with 0.3 µL DpnI at 37°C for at least 30 minutes and the entire mixture loaded in an agarose gel for electrophoresis. Fragments corresponding to the expected band sizes were excised and purified with a Zymoclean Gel DNA Recovery kit (Zymo Research) and then ligated with T4 ligase. A second round of PCR to add BsmBI cut sites and four-base overhangs was performed with 10 ng of each ligated vector as template. PCR reactions were digested with 0.3 µL of BsmBI and DpnI for 30 minutes at 37°C and 30 minutes at 55°C, then gel purified again. Golden Gate assembly was then performed, with all fragments and the DV diluted to 50 nM before use. 0.15 µL DV, 0.5 µL each gRNA part, 1 µl T4 DNA ligase (NEB), 1 µl 10X T4 ligase buffer (NEB) and 1.5 µl of BsmBI enzyme (NEB) were added to a 10 µl total reaction volume with ultrapure water. Assembly reactions were incubated for thirty cycles, with 5 min at 37°C and 5 min at 16°C, followed by a 5 min, 55°C final incubation step prior to transformation into TurboComp cells (NEB). Transformed cells were plated on LB + chloramphenicol and screened visually for colonies lacking GFP under blue light. Colonies lacking GFP upon visual inspection were picked, incubated at 37°C and their DNA isolated. Restriction digests were performed with 0.5µl BsaI (NEB) on 100 ng of DNA from each colony at 37°C for 1 hour and gel electrophoresis performed to verify

insert sizes. Of those colonies that yielded DNA bands of expected sizes, DNA from three colonies with each array (3, 6, 9, 12) were sent for additional validation by Sanger sequencing.

Construction of Single gRNA Vectors for Quantification of Repression

To construct gRNAs for targeting dCas9 to one of the three promoters driving the expression of a fluorescent protein, we used the pYTK050 part from the Yeast MoClo Toolkit[1]. pYTK050 contains a phenylalanine tRNA, HDV ribozyme, the Cas9/dCas9 scaffold (sgRNA) sequence and a SNR52 terminator. Complementary oligos were ordered (IDT) and annealed, with two additional 'T' nucleotides added before the spacer sequence to enable its ligation into pYTK050. Annealed oligos were then ligated into pYTK050 via BsmBI-mediated Golden Gate assembly. pYTK050 plasmids containing designed spacer sequences were genome-integrated at the URA3 locus.

Strains and Cell Cultures

All *Saccharomyces cerevisiae* strains generated in this study are derivatives of BY4741 (S288C MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0) and are listed in **Supplementary Table S14**. All experiments were performed in synthetic complete media with 2% (w/v) glucose (VWR), 0.67% (w/v) Yeast Nitrogen Base without amino acids (Sigma), 0.14% (w/v) Yeast Synthetic Drop-out Medium supplements without histidine, leucine, tryptophan, and uracil (Sigma), 20 mg/L histidine (Sigma), 100 mg/L leucine (Sigma), 20 mg/L tryptophan (Sigma), 20 mg/L uracil (Sigma).

Yeast colonies were inoculated into liquid Synthetic Complete media with all amino acids supplemented apart from the corresponding, auxotrophic amino acids (Uracil, Leucine and Histidine) and incubated in 96-well, 2.2 mL deep well plates at 30°C and 700 rpm.

Flow Cytometry

Cell fluorescence was measured by a LSRFortessa X-20 flow cytometer (BD Biosciences) using the following settings: FSC 330 V, SSC 250 V, mTagBFP2 - 405 nm line (450/50) 378 V, sfGFP - 488 nm line (530/30) 450 V, mRuby2 - 640 nm line (685/35) 590 V. Fluorescence data was collected from 10000 viable cells for each experiment and analyzed using FlowJo software.

SUPPORTING INFORMATION

Detailed protocol of the method to assemble gRNA arrays, including all DNA sequences, primers and parts used in this study, supporting data and additional methods.

ACKNOWLEDGMENTS

This work was supported by BBSRC grant (BB/R01602X/1) (to RLA), BBSRC grant BB/K019791/1 (to TE) and a Fulbright – Imperial College London Partnership award (to N.S.M).

DECLARATION OF INTERESTS

A patent application related to this work has been filed by the authors.

REFERENCES

(1) Lee, M. E., DeLoache, W. C., Cervantes, B., and Dueber, J. E. (2015) A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth. Biol. 4*, 975–986.
(2) Nissim, L., Perli, S. D., Fridkin, A., Perez-Pinera, P., and Lu, T. K. (2014) Multiplexed and Programmable Regulation of Gene Networks with an Integrated RNA and CRISPR/Cas Toolkit in Human Cells. *Mol. Cell 54*, 698–710.
(3) Gao, Y., and Zhao, Y. (2014) Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. *J. Integr. Plant Biol. 56*, 343–349.
(4) Nowak, C. M., Lawson, S., Zerez, M., and Bleris, L. (2016) Guide RNA engineering for versatile Cas9 functionality. *Nucleic Acids Res. 44*, 9555–9564.
(5) Ferreira, R., Skrekas, C., Nielsen, J., and David, F. (2018) Multiplexed CRISPR/Cas9 Genome Editing and Gene Regulation Using Csy4 in Saccharomyces cerevisiae. *ACS Synth. Biol. 7*, 10–15.
(6) Kurata, M., Wolf, N. K., Lahr, W. S., Weg, M. T., Kluesner, M. G., Lee, S., Hui, K., Shiraiwa, M., Webber, B. R., and Moriarity, B. S. (2018) Highly multiplexed genome engineering using CRISPR/Cas9 gRNA arrays. *PLoS One 13*, e0198714.
(7) Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K., and Doudna, J. A. (2010) Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science 329*, 1205–1210.
(8) Breunig, C. T., Durovic, T., Neuner, A. M., Baumann, V., Wiesbeck, M. F., Köferle, A., Götz, M., Ninkovic, J., and Stricker, S. H. (2018) One step generation of customizable gRNA vectors for multiplex CRISPR approaches through string assembly gRNA cloning (STAgR). *PLoS One 13*, e0196015.
(9) Xie, K., Minkenberg, B., and Yang, Y. (2015) Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl. Acad. Sci. 112*, 3570–3575.
(10) Sakuma, T., Nishikawa, A., Kume, S., Chayama, K., and Yamamoto, T. (2014) Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci. Rep. 4*, 5400.
(11) Vad-Nielsen, J., Lin, L., Bolund, L., Nielsen, A. L., and Luo, Y. (2016) Golden Gate Assembly of CRISPR gRNA expression array for simultaneously targeting multiple genes. *Cell. Mol. Life Sci. 73*, 4315–4325.
(12) Cress, B. F., Toparlak, O. D., Guleria, S., Lebovich, M., Stieglitz, J. T., Englaender, J. A., Jones, J. A., Linhardt, R. J., and Koffas, M. A. G. (2015) CRISPathBrick: Modular Combinatorial Assembly of Type II-A CRISPR Arrays for dCas9-Mediated Multiplex Transcriptional Repression in E. coli. *ACS Synth. Biol. 4*, 987–1000.
(13) Gander, M. W., Vrana, J. D., Voje, W. E., Carothers, J. M., and Klavins, E. (2017) Digital logic circuits in yeast with CRISPR-dCas9. *Nat. Commun. 8*, 1–11.
(14) Zhang, Y., Wang, J., Wang, Z., Zhang, Y., Shi, S., Nielsen, J., and Liu, Z. (2019) A gRNA-tRNA

array for CRISPR-Cas9 based rapid multiplexed genome editing in Saccharomyces cerevisiae. *Nat. Commun. 10*, 1–10.

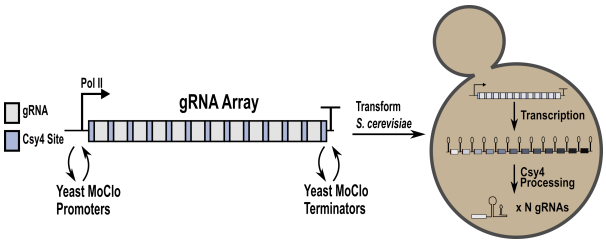
(15) Vladimir, P., Ong, J. L., Kucera, R. B., Langhorst, B. W., Bilotti, K., Pryor, J. M., Cantor, E. J., Canton, B., Knight, T. F., Evans, T. C., and Lohman, G. J. S. (2018) Comprehensive profiling of four base overhang ligation fidelity by T4 DNA Ligase and application to DNA assembly. *ACS Synth. Biol.* 8b00333.

(16) Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., Virgin, H. W., Listgarten, J., and Root, D. E. (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol. 34*, 184–191.

(17) Lian, J., Hamedirad, M., Hu, S., and Zhao, H. (2017) Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. *Nat. Commun. 8*, 1688.

(18) Jensen, E. D., Ferreira, R., Jakočiunas, T., Arsovska, D., Zhang, J., Ding, L., Smith, J. D., David, F., Nielsen, J., Jensen, M. K., and Keasling, J. D. (2017) Transcriptional reprogramming in yeast using dCas9 and combinatorial gRNA strategies. *Microb. Cell Fact. 16*, 1–16.

(19) Argueso, J. L., Westmoreland, J., Mieczkowski, P. A., Gawel, M., Petes, T. D., and Resnick, M. A. (2008) Double-strand breaks associated with repetitive DNA can reshape the genome. *Proc. Natl. Acad. Sci. 105*, 11845–11850.


Pol II
gRNA
Csy4 Site
gRNA Array
Yeast MoClo
Promoters
Yeast MoClo
Terminators
Transform
S. cerevisiae
Transcription
Csy4
Processing
x N gRNAs